Design of Nanny's Abnormal Behavior Recognition Bracelet Based on Human Activity Recognition(HAR) Deep Learning Model

Dengpeng Wang*, Yingfan Zhu, Yande Xiang, Ziyue Xaio
Tianjin University of Science and Technology, Tianjin, China

Email: 3264418107@qq.com

Abstract

This article designs a nanny abnormal behavior recognition bracelet. The bracelet is equipped with multiple sensors and a powerful control board, forming a complete nanny abnormal behavior recognition system, which realizes nanny abnormal behavior recognition and alarm in various environments. It uses a gyroscope to collect the three-axis acceleration information of the nanny, and uses the HAR model to infer the nanny's real-time behavior. When the abnormal behavior of the nanny is inferred, STM32 sends the GPS collected positioning information to the employer through ESP32 for timely alarm.

Keywords: Behavior recognition, Nanny, HAR model, Alarm

1. Introduction

With the increasing aging situation in China and the gradual implementation of the two-child and three-child policy, the number of elderly and young people in China has soared, and more and more elderly and children need to be taken care of, and the market scale of the nanny industry has expanded year by year. However, it is difficult to regulate the behavior of nannies, and the incidents of nannies abusing the caretakers are frequent. At present, employers can only monitor nannies by installing surveillance cameras, but there are large dead corners where nannies can still commit violent acts against the caretakers.In particular to dynamics analysis with moving bodies, the selection of the numerical integration method is crucial for the realization of the actual dynamics occurring in the real world. It cannot be solved by a simple way to chip the time step of the integration in the explicit numerical method. In the implicit numerical integration, it will help to refine the time step adaptively.

A nanny's abnormal behavior recognition bracelet based on HAR deep learning model was studied to solve such pain points. The nanny needs to wear the bracelet anytime and anywhere, and the employer can check the nanny's behavior in real time through the mobile phone. When the nanny's behavior to the caregiver is identified, the location of the nanny is directly sent to the employer so that the employer can timely understand the situation and report to the police.

The rest of this paper is organized as follows: The second part introduces the construction and deployment of HAR model. The third part introduces the hardware selection. The fourth part introduces the software design of the system. The fifth part summarizes the main content of this paper.

2. Construction and deployment of HAR model

2.1. *HAR Model Introduction*

Activity recognition has recently gained attention as a research topic because of the increasing availability of accelerometers in consumer products, like cell phones, and because of the many potential applications [1]. HAR model is an efficient human behavior recognition model, which aims to accurately monitor complex human behavior with limited hardware resources. The model uses a convolutional neural network (CNN [2]) architecture, a deep learning architecture that has achieved remarkable success in the field of image recognition and is particularly suitable for processing multidimensional time series data.

The reasoning principle of the model is that when the human body performs different actions, the three-axis acceleration curves of Roll, Pitch and Yaw are significantly different. Fig. 1 and Fig. 2 show the three-axis acceleration curves of the human body while sitting and jogging respectively, and it can be seen that the two curves are significantly different. The model can judge the specific behavior of the human body by the triaxial acceleration when performing different actions.

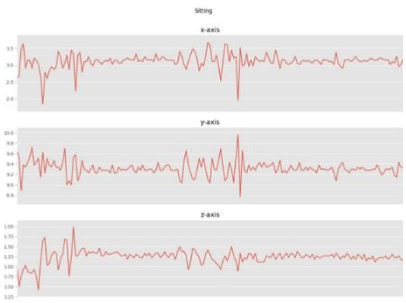


Fig.1 Three axis acceleration curve of human body during sitting still.

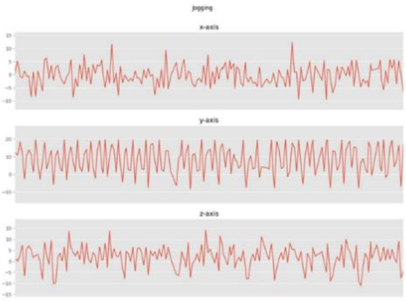


Fig.2 Triaxial Acceleration Curve of the Human Body During Jogging.

2.2. *Construction of HAR model*

In the process of model construction, we fully consider the advantages of convolutional neural network in time series data processing. The convolutional layer can effectively capture local patterns in the data, while the maximum pooling layer helps to reduce the sensitivity of the model to the input data and improve the robustness of the model. The fully connected layer is responsible for integrating these local information to form a holistic understanding of different behaviors. This hierarchical structure allows the model to better learn and understand complex behavioral features. The overall structure is shown in Fig. 3.

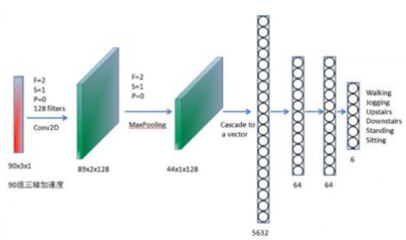


Fig. 3 Overall Network Structure of HAR Model

In the design of the network structure, the parameter configuration of each layer is paid attention to to ensure that the model can effectively capture the key features in the timing data. First, by setting the number of convolutional nuclei to 90, the size to 3*3, and the step size to 1, the local pattern of the input data is effectively extracted, which helps to capture subtle changes in the behavioral data. The activation function of the convolutional layer is ReLU [3] to introduce nonlinear characteristics and enhance the expressiveness of the model.

In the maximum pooling layer of the second layer, a 2*2 pooling kernel is adopted and its step size is set to 2 to reduce the spatial dimension of the data and the sensitivity of the model to the input data. This step helps to improve the generalization performance of the model so that it can better adapt to changes in different samples.

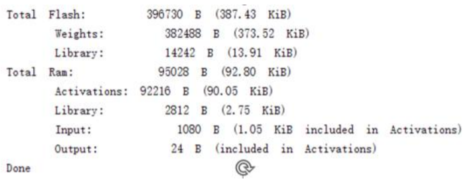
The design of the third convolutional layer continues to follow the above principles, the number of convolutional cores is 44, the size is 3*3, the step size is 1, the zero padding is 0, and the activation function is ReLU. The setting of this layer further deepens the feature extraction of time series data, and provides more informative input for the subsequent processing of the full connection layer.

The next three fully connected layers play a key role in the overall structure. The 5632 nodes of the fourth fully connected layer are responsible for integrating the local information extracted by the convolutional layer to form a holistic understanding of different behaviors. Two fully connected layers with a number of 64 nodes further improve the abstraction capability of the model, enabling it to better capture the abstract characteristics of the behavior.

The final output layer is a fully connected layer with six nodes, each corresponding to a behavior category (walking, jogging, going up, going down, standing, sitting, etc.). This design ensures that the model can accurately classify different behaviors, providing a reliable basis for real-time monitoring.

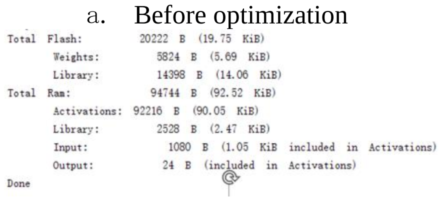
2.3. *Deployment of the HAR model*

In order to deploy the model on a small-capacity embedded device, we first upgraded the HAR model iteratively for several times, and successfully compressed the size of the HAR model from the initial 387.43KB to 19.75KB through operations such as pruning, quantization and optimization of the algorithm structure. The model size optimization effect is shown in Fig. 4.



```
Total  Flash:          396730  B  (387.43  KiB)
        Weights:          382488  B  (373.52  KiB)
        Library:          14242  B  (13.91  KiB)
Total  Ram:             95028  B  (92.80  KiB)
        Activations:  92216  B  (90.05  KiB)
        Library:          2812  B  (2.75  KiB)
        Input:              1080  B  (1.05  KiB  included  in  Activations)
        Output:            24  B  (included  in  Activations)
Done
```

a. Before optimization

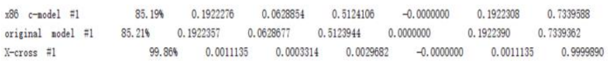


```
Total  Flash:          20222  B  (19.75  KiB)
        Weights:          5824  B  (5.69  KiB)
        Library:          14398  B  (14.06  KiB)
Total  Ram:             94744  B  (92.52  KiB)
        Activations:  92216  B  (90.05  KiB)
        Library:          2528  B  (2.47  KiB)
        Input:              1080  B  (1.05  KiB  included  in  Activations)
        Output:            24  B  (included  in  Activations)
Done
```

b. After optimization

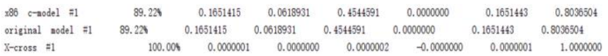
Fig. 4 Size comparison of HAR model before and after optimization

After lightweight and efficient processing, HAR model not only has a smaller storage footprint, but also can run on low-resource devices. In the training and design phase of the model, we employ a series of efficient algorithms and neural network structures, including advanced techniques such as convolutional neural networks (CNN) and recurrent neural networks (RNN [4]). These choices greatly reduce the training time and prediction time of the model to ensure both accuracy and timely reasoning in real-time applications. The improved model has achieved remarkable results in more than 100,000 test sets, and the accuracy rate of the model has increased from 85.19% to 89.22%, as shown in Fig. 5.



x86 c-model #1	85.19%	0.1922276	0.0628854	0.5124106	-0.0000000	0.1922308	0.7339588
original model #1	85.21%	0.1922357	0.0628677	0.5123944	0.0000000	0.1922390	0.7339362
X-cross #1	99.86%	0.0011135	0.0003314	0.0029682	-0.0000000	0.0011135	0.9999890

a. Before optimization



x86 c-model #1	89.22%	0.1651415	0.0618931	0.4544591	0.0000000	0.1651443	0.8036504
original model #1	89.22%	0.1651415	0.0618931	0.4544591	0.0000000	0.1651443	0.8036504
X-cross #1	100.00%	0.0000001	0.0000000	0.0000002	-0.0000000	0.0000001	1.0000000

b. After optimization

Fig. 5 Comparison of accuracy before and after HAR model optimization

The deployment of the model is realized by STM32CubeMX tool, and the Cube AI component in the software can compress the model according to different magnifications and convert it into C language, so as to facilitate the deployment of the model to the embedded development equipment.

3. Hardware selection

3.1. *Main control chip*

The main control chip adopts STM32F407VET6 developed by ST Company. The chip adopts Cortex-M4 core, and the peripherals are mounted with multi-channel I2C, SPI, CAN and other communication interfaces. It has rich functions and its main frequency reaches 168MHz. With 512KB Flash and 192KB RAM, the powerful computing power and sufficient memory are enough to meet the real-time requirements of nanny abnormal behavior recognition bracelet behavior reasoning. The outline design of the chip is shown in Fig. 6.

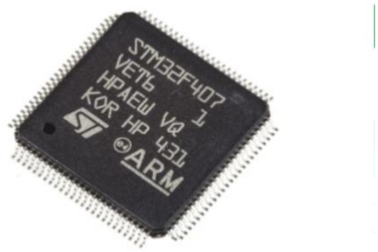


Fig. 6 STM32F407VET6 chip shape design

3.2. *WIFI module*

This design uses ESP32-WROOM-32WIFI module, its appearance is shown in Fig. 7. The ESP32-WROOM-32 is a powerful universal Wi-Fi+BT+BLE MCU module with powerful features and a wide range of application areas. It can be used in low-power sensor networks and high-performance tasks such as speech coding, audio streaming, and MP3 decoding.

Fig. 7 WIFI module appearance diagram

The module combines high performance computing and low power consumption, which is very suitable for iot project development application scenarios. It uses asynchronous serial port to communicate with STM32 to receive data, acts as a Wi-Fi access point or client, accesses Aliyun iot platform to achieve wireless network connection, and supports wireless communication between iot devices.

3.3. *Gyroscope*

The gyroscope in this design uses MPU6050 attitude sensor, and its chip appearance is shown in Fig. 8. The MPU6050 is a 6-axis attitude sensor that integrates a 3-axis MEMS gyroscope, a 3-axis MEMS accelerometer, and a scalable Digital Motion Processor (DMP). The range of the accelerometer can reach ±16g, and there is a configurable digital low-pass filter inside, and the configurable register can select the output data for low-pass filtering. The MPU6050 communicates with the STM32 via the IIC bus to transmit the nanny's three-axis acceleration.

Fig. 8 MPU6050 chip

3.4. *GPS Module*

The GPS module uses TAU1201 positioning module, and its module appearance is shown in Fig. 9. TAU1201 is a high-performance dual-band GNSS positioning module, equipped with BDA Beidou's CYNOSURE III GNSS SoC chip, the module supports the new generation of Beidou III signal system, and supports all civil navigation satellite systems in the world. These include BDS, GPS, GLONASS, Galileo, IRNSS, QZSS and SBAS (WAAS, EGNOS, GAGAN and MSAS). The TAU1201 integrates an efficient power management architecture to provide a high-precision, high-sensitivity, low-power solution for GNSS navigation applications.

Fig. 9 Appearance of the TAU1201 module

4. System software design

4.1. *Acquisition of triaxial acceleration*

The STM32 reads the contents of the relevant registers of the MPU6050 through the IIC communication protocol, a serial communication bus using a multi-master-slave architecture developed by Philips in the 1980s to allow motherboards, embedded systems, or mobile phones to connect low-speed peripherals. The IIC communication rate of STM32F407VET6 can reach up to 400KHz, which is enough to meet the requirements of HAR model inference.

Fig. 10 shows each meaning of the acceleration range register of the MPU6050 attitude sensor and the configuration guide. During configuration, the accelerometer range can be controlled by writing data to the register. Fig. 11 is an introduction of the acceleration read register of the MPU6050 attitude sensor. When the MPU6050 is used, the raw data of the three-axis acceleration is obtained through the 16-bit ADC value of the acceleration read register after the accelerometer RANGE is configured. The relationship between the three-axis acceleration (ACC) and the acceleration range (R) and the raw data (ADC) of the three-axis acceleration can be expressed as:

$$ACC = \frac{R * ADC}{32768} \tag{1}$$

Register (Hex)	Register (Decimal)	Bit7	Bit6	Bit5	Bit4	Bit3	Bit2	Bit1	Bit0
1C	28	XA_ST	YA_ST	ZA_ST	AFS_SEL[1:0]		ACCEL_HPF[2:0]		

Description:

This register is used to trigger accelerometer self test and configure the accelerometer full scale range. This register also configures the Digital High Pass Filter (DHPF).

Accelerometer self-test permits users to test the mechanical and electrical portions of the accelerometer. The self-test for each accelerometer axis can be activated by controlling the *XA_ST*, *YA_ST*, and *ZA_ST* bits of this register. Self-test for each axis may be performed independently or all at the same time.

When self-test is activated, the on-board electronics will actuate the appropriate sensor. This actuation simulates an external force. The actuated sensor, in turn, will produce a corresponding output signal. The output signal is used to observe the self-test response.

The self-test response is defined as follows:

Self-test response = Sensor output with self-test enabled – Sensor output without self-test enabled

The self-test limits for each accelerometer axis is provided in the electrical characteristics tables of the MPU-6000/MPU-6050 Product Specification document. When the value of the self-test response is within the min/max limits of the product specification, the part has passed self test. When the self-test response exceeds the min/max values specified in the document, the part is deemed to have failed self-test.

AFS_SEL selects the full scale range of the accelerometer outputs according to the following table.

AFS_SEL	Full Scale Range
0	± 2*g*
1	± 4*g*
2	± 8*g*
3	± 16*g*

Fig. 10 MPU6050 aceleration configuration register

Register (Hex)	Register (Decimal)	Bit7	Bit6	Bit5	Bit4	Bit3	Bit2	Bit1	Bit0
3B	59	ACCEL_XOUT[15:8]							
3C	60	ACCEL_XOUT[7:0]							
3D	61	ACCEL_YOUT[15:8]							
3E	62	ACCEL_YOUT[7:0]							
3F	63	ACCEL_ZOUT[15:8]							
40	64	ACCEL_ZOUT[7:0]							

Description:

These registers store the most recent accelerometer measurements.

Accelerometer measurements are written to these registers at the Sample Rate as defined in Register 25.

The accelerometer measurement registers, along with the temperature measurement registers, gyroscope measurement registers, and external sensor data registers, are composed of two sets of registers: an internal register set and a user-facing read register set.

The data within the accelerometer sensors' internal register set is always updated at the Sample Rate. Meanwhile, the user-facing read register set duplicates the internal register set's data values whenever the serial interface is idle. This guarantees that a burst read of sensor registers will read measurements from the same sampling instant. Note that if burst reads are not used, the user is responsible for ensuring a set of single byte reads correspond to a single sampling instant by checking the Data Ready interrupt.

Each 16-bit accelerometer measurement has a full scale defined in *ACCEL_FS* (Register 28). For each full scale setting, the accelerometers' sensitivity per LSB in *ACCEL_xOUT* is shown in the table below.

AFS_SEL	Full Scale Range	LSB Sensitivity
0	±2*g*	16384 LSB/mg
1	±4*g*	8192 LSB/mg
2	±8*g*	4096 LSB/mg
3	±16*g*	2048 LSB/mg

Fig. 11 MPU6050 acceleration read register

4.2. *Nanny location acquisition*

The nanny location is obtained through the TAU1201 module. TAU1201 module can capture the signal of the satellite to be tested according to a certain height cutoff Angle of the satellite, track the operation of these satellites, transform, amplify and process the received GPS signal, so as to measure the propagation time of the GPS signal from the satellite to the receiver antenna, and interpret the navigation message sent by the GPS satellite. When the TAU1201 is finished interpreting, it communicates with STM32 through its own serial port to send the nanny location information to STM32.

4.3. *Behavioral inference data and location information upload*

Data upload is completed through the ESP32-Wroom-32wifi module. When the nanny's behavior is deduced on STM32, the HAR model inference result will be sent to ESP32 through STM32's own serial port. After the ESP32 makes the judgment of dangerous behavior, the judgment result will be sent to the employer's mobile phone server through MQTT protocol. When it is deduced that the nanny is in danger, the nanny's real-time location information is uploaded to the employer's mobile phone so that the employer can take the next step.

5. Conclusion

The nanny abnormal behavior recognition bracelet, to a large extent, solves the problems of traditional surveillance cameras, such as large monitoring dead Angle and not easy to carry, so that the nanny's behavior can be better supervised and standardized, has a good role in promoting the development of the nanny industry, and indirectly protects the life and property safety of the caretakers.

References

1. Jennifer R. Kwapisz, Gary M. Weiss, Samuel A. Moore, Activity Recognition using Cell Phone Accelerometers, ACM SIGKDD Explorations Newsletter.Volume 12,Issue 2.2011: pp.74-82.
2. Wang Qinghua;Li Ziwei;Zhang Shuqi;Chi Nan;Dai Qionghai,A versatile Wavelet-Enhanced CNN-Transformer for improved fluorescence microscopy image restoration,Neural Networks.Volume 170,Issue. 2024:pp.227-241.
3. He Zihao;Shu Qianyu;Wang Yinghua;Wen Jinming,A ReLU-based hard-thresholding algorithm for non-negative sparse signal recovery,Signal Processing.Volume 215,Issue.2024.
4. Xu Ziming;Leung Juliana Y,A novel formulation of RNN-based neural network with real-time updating – An application for dynamic hydraulic fractured shale gas production forecasting,Geoenergy Science and Engineering.Volume 233,Issue.2024.

Authors Introduction

Mr. DePeng Wang

He is studying at the School of Electronic Information and Automation at Tianjin University of Science and Technology, China.

Ms. YingFan Zhu

She is currently pursuing her undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Mr. YanDe Xiang

He is currently pursuing his undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.

Mr. ZiYue Xiao

He is currently pursuing his undergraduate degree at the School of Electronic Information and Automation, Tianjin University of Science and Technology.